

C1 Controls and Latin-1 Supplement

Range: 0080–00FF

DRAFT The Unicode Standard, Version 16.0 BETA REVIEW

This file contains an excerpt from the character code tables and list of character names for *DRAFT The Unicode Standard, Version 16.0 BETA REVIEW*

This file may be changed at any time without notice to reflect errata, or other updates to the Unicode Standard. See <https://www.unicode.org/errata/> for an up-to-date list of errata.

See <https://www.unicode.org/charts/> for access to a complete list of the latest character code charts. See <https://www.unicode.org/charts/PDF/Unicode-15.1/> for charts showing only the characters added in Unicode 15.1. See <https://www.unicode.org/Public/15.1.0/charts/> for a complete archived file of character code charts for Unicode 15.1. See <https://www.unicode.org/charts/About.html#Conventions> for conventions used in these code charts, and other general information.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 15.1 but do not provide all the information needed to fully support individual scripts using the Unicode Standard. For a complete understanding of the use of the characters contained in this file, please consult the appropriate sections of The Unicode Standard, Version 15.1, online at <https://www.unicode.org/versions/Unicode15.1.0/>, as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #29, #31, #34, #38, #41, #42, #44, #45, and #50, the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

See <https://www.unicode.org/ucd/> and <http://www.unicode.org/reports/>

A thorough understanding of the information contained in these additional sources is required for a successful implementation.

Copying characters from the character code tables or list of character names is not recommended, because for production reasons the PDF files for the code charts cannot guarantee that the correct character codes will always be copied.

Fonts

The shapes of the reference glyphs used in these code charts are not prescriptive. Considerable variation is to be expected in actual fonts. The particular fonts used in these charts were provided to the Unicode Consortium by a number of different font designers, who own the rights to the fonts.

See <https://www.unicode.org/charts/fonts.html> for a list.

Terms of Use

You may freely use these code charts for personal or internal business uses only. You may not incorporate them either wholly or in part into any product or publication, or otherwise distribute them without express written permission from the Unicode Consortium. However, you may provide links to these charts.

The fonts and font data used in production of these code charts may NOT be extracted, or used in any other way in any product or publication, without permission or license granted by the typeface owner(s).

The Unicode Consortium is not liable for errors or omissions in this file or the standard itself. Information on characters added to the Unicode Standard since the publication of the most recent version of the Unicode Standard, as well as on characters currently being considered for addition to the Unicode Standard can be found on the Unicode web site.

See <https://www.unicode.org/pending/pending.html> and <http://www.unicode.org/alloc/Pipeline.html>.

See <https://www.unicode.org/charts/About.html> for more information concerning the conventions and symbols used in these code charts.

Copyright © 1991–2023 Unicode, Inc. All rights reserved. See <https://www.unicode.org/copyright.html>

	008	009	00A	00B	00C	00D	00E	00F
0	XXX 0080	DCS 0090	NB SP 00A0	◊ 00B0	À 00C0	Ð 00D0	à 00E0	ð 00F0
1	XXX 0081	PU1 0091	¡ 00A1	± 00B1	Á 00C1	Ñ 00D1	á 00E1	ñ 00F1
2	BPH 0082	PU2 0092	¢ 00A2	² 00B2	Â 00C2	Ò 00D2	â 00E2	ò 00F2
3	NBH 0083	STS 0093	£ 00A3	³ 00B3	Ã 00C3	Ó 00D3	ã 00E3	ó 00F3
4	IND 0084	CCH 0094	¤ 00A4	´ 00B4	Ä 00C4	Ô 00D4	ä 00E4	ô 00F4
5	NEL 0085	MW 0095	¥ 00A5	µ 00B5	Å 00C5	Õ 00D5	å 00E5	õ 00F5
6	SSA 0086	SPA 0096	¦ 00A6	¶ 00B6	Æ 00C6	Ö 00D6	æ 00E6	ö 00F6
7	ESA 0087	EPA 0097	§ 00A7	· 00B7	Ç 00C7	× 00D7	ç 00E7	÷ 00F7
8	HTS 0088	SOS 0098	¨ 00A8	¸ 00B8	È 00C8	Ø 00D8	è 00E8	ø 00F8
9	HTJ 0089	XXX 0099	© 00A9	¹ 00B9	É 00C9	Ù 00D9	é 00E9	ù 00F9
A	VTS 008A	SCI 009A	ª 00AA	º 00BA	Ê 00CA	Ú 00DA	ê 00EA	ú 00FA
B	PLD 008B	CSI 009B	« 00AB	» 00BB	Ë 00CB	Û 00DB	ë 00EB	û 00FB
C	PLU 008C	ST 009C	¬ 00AC	¼ 00BC	Ì 00CC	Ü 00DC	ì 00EC	ü 00FC
D	RI 008D	OSC 009D	SHY 00AD	½ 00BD	Í 00CD	Ý 00DD	í 00ED	ý 00FD
E	SS2 008E	PM 009E	® 00AE	¾ 00BE	Î 00CE	Þ 00DE	î 00EE	þ 00FE
F	SS3 008F	APC 009F	¯ 00AF	¿ 00BF	Ï 00CF	ß 00DF	ï 00EF	ÿ 00FF

C1 controls

Alias names are those for ISO/IEC 6429:1992.

0080		<control>
0081		<control>
0082		<control> = BREAK PERMITTED HERE → 200B zero width space
0083		<control> = NO BREAK HERE → 2060 word joiner
0084		<control> • formerly known as INDEX
0085		<control> = NEXT LINE (NEL)
0086		<control> = START OF SELECTED AREA
0087		<control> = END OF SELECTED AREA
0088		<control> = CHARACTER TABULATION SET
0089		<control> = CHARACTER TABULATION WITH JUSTIFICATION
008A		<control> = LINE TABULATION SET
008B		<control> = PARTIAL LINE FORWARD
008C		<control> = PARTIAL LINE BACKWARD
008D		<control> = REVERSE LINE FEED
008E		<control> = SINGLE SHIFT TWO
008F		<control> = SINGLE SHIFT THREE
0090		<control> = DEVICE CONTROL STRING
0091		<control> = PRIVATE USE ONE
0092		<control> = PRIVATE USE TWO
0093		<control> = SET TRANSMIT STATE
0094		<control> = CANCEL CHARACTER
0095		<control> = MESSAGE WAITING
0096		<control> = START OF GUARDED AREA
0097		<control> = END OF GUARDED AREA
0098		<control> = START OF STRING
0099		<control>
009A		<control> = SINGLE CHARACTER INTRODUCER
009B		<control> = CONTROL SEQUENCE INTRODUCER
009C		<control> = STRING TERMINATOR
009D		<control> = OPERATING SYSTEM COMMAND
009E		<control> = PRIVACY MESSAGE

009F <control>
= APPLICATION PROGRAM COMMAND

Latin-1 punctuation and symbols

Based on ISO/IEC 8859-1 (aka Latin-1) from here.

00A0		NO-BREAK SPACE • commonly abbreviated as NBSP → 0020 space → 2007 figure space → 202F narrow no-break space → 2060 word joiner → 237D shouldered open box → FEFF zero width no-break space ≈ <noBreak> 0020
00A1		INVERTED EXCLAMATION MARK • Spanish, Asturian, Galician • used in extended IPA to represent sublingual lower alveolar percussive → 0021 exclamation mark → A71E modifier letter raised inverted exclamation mark
00A2		CENT SIGN → 023C latin small letter c with stroke → 20A5 mill sign → 20B5 cedi sign → 1D135 musical symbol cut time
00A3		POUND SIGN = pound sterling, Irish punt, lira, etc. • not to be confused with the unit of weight • glyph may have one or two crossbars, in distinction from 20A4 → 20A4 lira sign → 20BA turkish lira sign → 10192 roman semuncia sign
00A4		CURRENCY SIGN • generic or unknown currency • other currency symbol characters start at 20A0 → 0024 dollar sign
00A5		YEN SIGN = yuan sign (renminbi) • glyph may have one or two crossbars, but the official symbol in Japan and China has two • in native context, specific ideographs may be used for units of these currencies → 024E latin capital letter y with stroke → 04B0 cyrillic capital letter straight u with stroke → 5143 元 → 5186 円
00A6		BROKEN BAR = broken vertical bar (1.0) = parted rule (in typography) → 007C vertical line
00A7		SECTION SIGN • used in referencing numbered sections or paragraphs → 2E39 top half section sign
00A8		DIAERESIS • this is a spacing character → 0308 combining diaeresis ≈ 0020 0308

00A9	©	COPYRIGHT SIGN → 2117 © sound recording copyright → 24B8 © circled latin capital letter c → 1F12F © copyleft symbol → 1F1AD © mask work symbol	00B4	´	ACUTE ACCENT • this is a spacing character → 0060 ` grave accent → 02B9 ´ modifier letter prime → 02CA ´ modifier letter acute accent → 02DD ¨ double acute accent → 0301 ó combining acute accent → 0384 ´ greek tonos → 1FFD ´ greek oxia → 2032 ´ prime ≈ 0020 [SP] 0301 ó
00AA	ª	FEMININE ORDINAL INDICATOR • Spanish → 1D43 ª modifier letter small a ≈ <super> 0061 a	00B5	µ	MICRO SIGN ≈ 03BC µ greek small letter mu
00AB	«	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK = left guillemet = chevrons (in typography) • usually opening, sometimes closing → 226A « much less-than → 300A ≪ left double angle bracket	00B6	¶	PILCROW SIGN = paragraph sign • section sign in some European usage → 204B ¶ reversed pilcrow sign → 2761 ¶ curved stem paragraph sign ornament → 2E3F ¶ capitulum
00AC	¬	NOT SIGN = angled dash (in typography) → 2310 ¬ reversed not sign → 2319 ⌵ turned not sign	00B7	·	MIDDLE DOT = midpoint (in typography) = Georgian comma = Greek middle dot (ano teleia) • also used as a raised decimal point or to denote multiplication; for multiplication 22C5 · is preferred • used as a vowel length mark (part of words) in many Amerindian orthographies → 002E . full stop → 02D9 ¨ dot above → 0387 · greek ano teleia → 2022 • bullet → 2024 . one dot leader → 2027 · hyphenation point → 2219 • bullet operator → 22C5 · dot operator → 2E31 · word separator middle dot → 2E33 · raised dot → 30FB • katakana middle dot → A78F • latin letter sinological dot
00AD	⸀	SOFT HYPHEN = discretionary hyphen • commonly abbreviated as SHY • an invisible format character indicating a possible hyphenation location	00B8	¸	CEDILLA • this is a spacing character • other spacing accent characters: 02D8 ¸—02DB ¸ → 0327 ¸ combining cedilla ≈ 0020 [SP] 0327 ¸
00AE	®	REGISTERED SIGN = registered trade mark sign (1.0) → 2122 ™ trade mark sign → 24C7 ® circled latin capital letter r → 1F16B ™ raised md sign → 1F16C ™ raised mr sign	00B9	¹	SUPERSCRRIPT ONE → 00B2 ² superscript two → 00B3 ³ superscript three → 2070 ⁰ superscript zero ≈ <super> 0031 1
00AF	—	MACRON = overline, APL overbar • this is a spacing character • unlike 203E ¯, adjacent instances of this character do not join → 02C9 ¯ modifier letter macron → 0304 ¯ combining macron → 0305 ¯ combining overline → 203E ¯ overline → FE26 ¯ combining conjoining macron ≈ 0020 [SP] 0304 ¯	00BA	º	MASCULINE ORDINAL INDICATOR • Spanish → 1D52 º modifier letter small o → 2070 ⁰ superscript zero → 2116 № numero sign ≈ <super> 006F o
00B0	°	DEGREE SIGN • this is a spacing character → 02DA ° ring above → 030A ° combining ring above → 2070 ⁰ superscript zero → 2103 °C degree celsius → 2109 °F degree fahrenheit → 2218 ° ring operator	00BB	»	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK = right guillemet • usually closing, sometimes opening → 226B » much greater-than → 300B ≫ right double angle bracket
00B1	±	PLUS-MINUS SIGN = biquintile (astrological aspect: 144°) → 2213 ± minus-or-plus sign			
00B2	²	SUPERSCRRIPT TWO = squared • other superscript digit characters: 2070 ⁰—2079 ⁹ → 00B9 ¹ superscript one ≈ <super> 0032 2			
00B3	³	SUPERSCRRIPT THREE = cubed → 00B9 ¹ superscript one ≈ <super> 0033 3			

Vulgar fractions

The fraction bar for these may be rendered horizontally or at a slant. For other fraction characters, see 2150-215E.

00BC	¼	VULGAR FRACTION ONE QUARTER	≈ <fraction> 0031 1 2044/ 0034 4
00BD	½	VULGAR FRACTION ONE HALF	≈ <fraction> 0031 1 2044/ 0032 2
00BE	¾	VULGAR FRACTION THREE QUARTERS	≈ <fraction> 0033 3 2044/ 0034 4

Punctuation

00BF	¿	INVERTED QUESTION MARK	= turned question mark
			• Spanish
			→ 003F ? question mark
			→ 2E18 ¿ inverted interrobang
			→ 2E2E ¿ reversed question mark

Uppercase letters

00C0	À	LATIN CAPITAL LETTER A WITH GRAVE	≡ 0041 A 0300 ò
00C1	Á	LATIN CAPITAL LETTER A WITH ACUTE	≡ 0041 A 0301 ó
00C2	Â	LATIN CAPITAL LETTER A WITH CIRCUMFLEX	≡ 0041 A 0302 ô
00C3	Ã	LATIN CAPITAL LETTER A WITH TILDE	≡ 0041 A 0303 õ
00C4	Ä	LATIN CAPITAL LETTER A WITH DIAERESIS	≡ 0041 A 0308 ö
00C5	Å	LATIN CAPITAL LETTER A WITH RING ABOVE	→ 212B Å angstrom sign
			≡ 0041 A 030A ø
00C6	Æ	LATIN CAPITAL LETTER AE	= latin capital ligature ae (1.1)
			→ 04D4 А cyrillic capital ligature a ie
00C7	Ç	LATIN CAPITAL LETTER C WITH CEDILLA	≡ 0043 C 0327 ç
00C8	È	LATIN CAPITAL LETTER E WITH GRAVE	≡ 0045 E 0300 è
00C9	É	LATIN CAPITAL LETTER E WITH ACUTE	≡ 0045 E 0301 é
00CA	Ê	LATIN CAPITAL LETTER E WITH CIRCUMFLEX	≡ 0045 E 0302 ê
00CB	Ë	LATIN CAPITAL LETTER E WITH DIAERESIS	≡ 0045 E 0308 ë
00CC	Ì	LATIN CAPITAL LETTER I WITH GRAVE	≡ 0049 I 0300 ì
00CD	Í	LATIN CAPITAL LETTER I WITH ACUTE	≡ 0049 I 0301 í
00CE	Î	LATIN CAPITAL LETTER I WITH CIRCUMFLEX	≡ 0049 I 0302 î
00CF	Ï	LATIN CAPITAL LETTER I WITH DIAERESIS	≡ 0049 I 0308 ï
00D0	Ð	LATIN CAPITAL LETTER ETH	• lowercase is 00F0 ð
			→ 0110 Ð latin capital letter d with stroke
			→ 0189 Ð latin capital letter african d
			→ A7C7 Ð latin capital letter d with short stroke overlay
00D1	Ñ	LATIN CAPITAL LETTER N WITH TILDE	≡ 004E N 0303 ñ
00D2	Ò	LATIN CAPITAL LETTER O WITH GRAVE	≡ 004F O 0300 ò

00D3	Ó	LATIN CAPITAL LETTER O WITH ACUTE	≡ 004F O 0301 ó
00D4	Ô	LATIN CAPITAL LETTER O WITH CIRCUMFLEX	≡ 004F O 0302 ô
00D5	Õ	LATIN CAPITAL LETTER O WITH TILDE	≡ 004F O 0303 õ
00D6	Ö	LATIN CAPITAL LETTER O WITH DIAERESIS	≡ 004F O 0308 ö

Mathematical operator

00D7	×	MULTIPLICATION SIGN	= z notation Cartesian product
			→ 0078 x latin small letter x
			→ 2613 X saltire
			→ 2715 X multiplication x
			→ 2718 X heavy ballot x
			→ 274C X cross mark
			→ 2A2F x vector or cross product
			→ 1F5D9 X cancellation x
			→ 1F7A9 X light saltire

Uppercase letters

00D8	Ø	LATIN CAPITAL LETTER O WITH STROKE	= o slash
			→ 019F Ø latin capital letter o with middle tilde
			→ 2205 Ø empty set
			→ A74A Ø latin capital letter o with long stroke overlay
00D9	Ù	LATIN CAPITAL LETTER U WITH GRAVE	≡ 0055 U 0300 ù
00DA	Ú	LATIN CAPITAL LETTER U WITH ACUTE	≡ 0055 U 0301 ú
00DB	Û	LATIN CAPITAL LETTER U WITH CIRCUMFLEX	≡ 0055 U 0302 û
00DC	Ü	LATIN CAPITAL LETTER U WITH DIAERESIS	≡ 0055 U 0308 ü
00DD	Ý	LATIN CAPITAL LETTER Y WITH ACUTE	≡ 0059 Y 0301 ý
00DE	Þ	LATIN CAPITAL LETTER THORN	

Lowercase letters

00DF	ß	LATIN SMALL LETTER SHARP S	= Eszett
			• German
			• not used in Swiss High German
			• uppercase is "SS" (standard case mapping), alternatively 1E9E Œ
			• typographically the glyph for this character can be based on a ligature of 017F f with either 0073 s or with an old-style glyph for 007A z (the latter similar in appearance to 0292 Ʒ). Both forms exist interchangeably today.
			→ 017F f latin small letter long s
			→ 0292 Ʒ latin small letter ezh
			→ 03B2 β greek small letter beta
			→ 1E9E Œ latin capital letter sharp s
			→ A7B5 β latin small letter beta
			→ A7D7 f latin small letter middle scots s
00E0	à	LATIN SMALL LETTER A WITH GRAVE	≡ 0061 a 0300 à
00E1	á	LATIN SMALL LETTER A WITH ACUTE	≡ 0061 a 0301 á
00E2	â	LATIN SMALL LETTER A WITH CIRCUMFLEX	≡ 0061 a 0302 â

00E3	ã	LATIN SMALL LETTER A WITH TILDE
		• Portuguese
		≡ 0061 a 0303 ǎ
00E4	ä	LATIN SMALL LETTER A WITH DIAERESIS
		≡ 0061 a 0308 ӓ
00E5	å	LATIN SMALL LETTER A WITH RING ABOVE
		• Danish, Norwegian, Swedish, Walloon
		≡ 0061 a 030A Ȧ
00E6	æ	LATIN SMALL LETTER AE
		= latin small ligature ae (1.1)
		= ash (from Old English æsc)
		• Danish, Norwegian, Icelandic, Faroese, Old English, French, IPA
		→ 0153 œ latin small ligature oe
		→ 04D5 æ cyrillic small ligature a ie
		→ 1D02 æ latin small letter turned ae
		→ AB31 æ latin small letter a reversed-schwa
00E7	ç	LATIN SMALL LETTER C WITH CEDILLA
		≡ 0063 c 0327 ̸
00E8	è	LATIN SMALL LETTER E WITH GRAVE
		≡ 0065 e 0300 ̀
00E9	é	LATIN SMALL LETTER E WITH ACUTE
		≡ 0065 e 0301 ́
00EA	ê	LATIN SMALL LETTER E WITH CIRCUMFLEX
		≡ 0065 e 0302 ˆ
00EB	ë	LATIN SMALL LETTER E WITH DIAERESIS
		≡ 0065 e 0308 ӓ
00EC	ì	LATIN SMALL LETTER I WITH GRAVE
		• Italian, Malagasy
		≡ 0069 i 0300 ̀
00ED	í	LATIN SMALL LETTER I WITH ACUTE
		≡ 0069 i 0301 ́
00EE	î	LATIN SMALL LETTER I WITH CIRCUMFLEX
		≡ 0069 i 0302 ˆ
00EF	ï	LATIN SMALL LETTER I WITH DIAERESIS
		≡ 0069 i 0308 ӓ
00F0	ð	LATIN SMALL LETTER ETH
		• Icelandic, Faroese, Old English, IPA
		• uppercase is 00D0 Ð
		→ 0111 đ latin small letter d with stroke
		→ 03B4 δ greek small letter delta
		→ 1E9F δ latin small letter delta
		→ 2202 ∂ partial differential
		→ A77A ð latin small letter insular d
00F1	ñ	LATIN SMALL LETTER N WITH TILDE
		≡ 006E n 0303 ̃
00F2	ò	LATIN SMALL LETTER O WITH GRAVE
		≡ 006F o 0300 ̀
00F3	ó	LATIN SMALL LETTER O WITH ACUTE
		≡ 006F o 0301 ́
00F4	ô	LATIN SMALL LETTER O WITH CIRCUMFLEX
		≡ 006F o 0302 ˆ
00F5	õ	LATIN SMALL LETTER O WITH TILDE
		• Portuguese, Estonian, Livonian
		≡ 006F o 0303 ̃
00F6	ö	LATIN SMALL LETTER O WITH DIAERESIS
		≡ 006F o 0308 ӓ

Mathematical operator

00F7	÷	DIVISION SIGN
		= obelus
		• occasionally used as an alternate, more visually distinct shape for a hyphen in some contexts
		• historically used as a punctuation mark to denote questionable passages in manuscripts
		→ 070B ⁂ syriac harklean obelus
		→ 2052 ⁄ commercial minus sign
		→ 2215 / division slash
		→ 2223 divides
		→ 2236 : ratio
		→ 2797 ÷ heavy division sign
		→ 27CC ⁝ long division
		→ 2E13 ⁞ dotted obelos

Lowercase letters

00F8	ø	LATIN SMALL LETTER O WITH STROKE
		= o slash
		• Danish, Norwegian, Faroese, IPA
		→ 0275 ø latin small letter barred o
		→ 2300 ø diameter sign
		→ A74B ø latin small letter o with long stroke overlay
		→ A7C1 ȯ latin small letter old polish o
		→ AB3E ø latin small letter blackletter o with stroke
00F9	ù	LATIN SMALL LETTER U WITH GRAVE
		• French, Italian
		≡ 0075 u 0300 ̀
00FA	ú	LATIN SMALL LETTER U WITH ACUTE
		≡ 0075 u 0301 ́
00FB	û	LATIN SMALL LETTER U WITH CIRCUMFLEX
		≡ 0075 u 0302 ˆ
00FC	ü	LATIN SMALL LETTER U WITH DIAERESIS
		≡ 0075 u 0308 ӓ
00FD	ý	LATIN SMALL LETTER Y WITH ACUTE
		• Czech, Slovak, Icelandic, Faroese, Welsh, Malagasy
		≡ 0079 y 0301 ́
00FE	þ	LATIN SMALL LETTER THORN
		• Icelandic, Old English, phonetics
		• Runic letter borrowed into Latin script
		• replaced by "th" in modern English
		→ 03F8 þ greek small letter sho
		→ 16A6 þ runic letter thurisaz thurs thorn
		→ A7D3 þ latin small letter double thorn
00FF	ÿ	LATIN SMALL LETTER Y WITH DIAERESIS
		• French, medieval Hungarian orthography
		• uppercase is 0178 Ÿ
		≡ 0079 y 0308 ӓ